How private is private information?

Edinburgh School of Economics

Discussion Paper Series

Number 237

How private is private information? The ability to spot deception in an economic game.

Michèle Belot
(University of Edinburgh)

Jeroen van de Ven
(Amsterdam Center for Law and Economics,
University of Amsterdam, and Tinbergen Institute)

Date
December 2013

Published by
School of Economics
University of Edinburgh
30 -31 Buccleuch Place
Edinburgh EH8 9JT
+44 (0)131 650 8361
http://edin.ac/16ja6A6

THE UNIVERSITY *of* EDINBURGH

How private is private information?

The ability to spot deception in an economic game.

Michèle Belot and Jeroen van de Ven*

This version: December 2013.

Abstract: We provide experimental evidence on the ability to detect deceit in a buyer-seller game with asymmetric information. Sellers have private information about the buyer's valuation of a good and sometimes have incentives to mislead buyers. We examine if buyers can spot deception in face-to-face encounters. We vary (1) whether or not the buyer can interrogate the seller, and (2) the contextual richness of the situation. We find that the buyers' prediction accuracy is above chance levels, and that interrogation and contextual richness are important factors determining the accuracy. These results show that there are circumstances in which part of the information asymmetry is eliminated by people's ability to spot deception.

Keywords: Deception, lie detection, asymmetric information, face-to-face interaction, experiment.

JEL codes: C91, D82, K4

* Belot: School of Economics, University of Edinburgh, michele.belot@ed.ac.uk. Van de Ven: Amsterdam Center for Law and Economics, University of Amsterdam, and Tinbergen Institute, j.vandeven@uva.nl. We are very grateful to seminar attendants in Lyon, Toulouse, Aix-Marseille, East Anglia, Amsterdam (Tinbergen Institute), Munich, Mainz, the 2013 EWEBE Meeting in Frankfurt, the ESA meetings in New York, the TIBER symposium, the MBEES workshop, and the ACLE spring meeting. Special thanks go to Arthur Schram, Gönül Doğan, and Jo Seldeslachts for their valuable suggestions and help with running the experiments, and to Christoph Engel for his careful thoughts as a discussant at the ACLE workshop.

1. Introduction

The existence and consequences of asymmetric information has a very prominent place in economic analysis. The limited transmission of information in those situations often causes inefficiencies. Two assumptions that are typically made in the literature are responsible for this. First, people with private information will not disclose their information unless it is in their narrow self-interest to do so. This implies that they are always willing to lie.[1] The second assumption is that the less informed people are not able to infer any private information beyond what they are able to infer from the equilibrium properties. If sellers always try to persuade buyers to purchase the most expensive item, or defendants always plead guilty, the less informed party cannot discriminate between honest or dishonest statements. In reality, there can be signals that (imperfectly) separate honest from dishonest people, and it is plausible that in some contexts people are able to process such signals and identify attempts of deception.

Despite its relevance for a wide range of economic situations, there is hardly any evidence in economics on the existence or relevance of signals that help people to detect deceit. In an intriguing recent experiment, Wang et al. (2010), show that in a sender-receiver game with private information, the senders' pupils dilate more when their deception is larger in magnitude. In fact, they find that if the receiver could use the information on pupil dilation to predict the sender's message, she could increase her payoff by 16–21 percent. There is further evidence that other verbal and non-verbal cues may be systematically associated with deception (Ekman et al., 1988; ten Brinke et al., 2012, Ekman, 1988). Thus, one key question is to what extent people are able to *detect* these signals and spot deception? Recent work provides encouraging evidence in that respect (e.g., Belot et al., 2012), but only very little is known about the scope and conditions under which people can detect deceit.

In this study, we use data from a laboratory experiment to estimate people's ability to detect lies in face-to-face encounters with free format communication.[2] Face-to-face interaction remains

[1] Several studies show that people are sometimes averse to lying (e.g., Gneezy, 2005; Charness and Dufwenberg, 2006a) and analyze the consequences this may have (e.g., Kartik, 2009; Kartik et al., 2007).

[2] As a shorthand we will adopt the terminology of lie detection, but strictly speaking it would be more appropriate to speak about *information transmission* in our setting. Information transmission is a broader concept that includes the detection of lies.

one of the most popular tools of communication. For example, in a recent global survey of 2300 Harvard Business Review subscribers[3], respondents said that face-to-face meetings are key to building long-term relationships (95%), negotiating contracts (89%), meeting new clients (79%), and understanding and listening to important customers (69%).

Our setting resembles an economically meaningful relationship. Two participants are matched, one in the role of seller and the other in the role of buyer. The buyer can purchase either one of two products, but only the seller is privately informed about which of the two products matches the buyer's interests. One of those products yields the seller a higher profit, and it is common knowledge which product that is. The experiment is set up such that the interests of the seller and buyer are sometimes aligned, so that the most profitable product matches the preferences of the buyer, while at other times there is a conflict of interest. Sellers are given the opportunity to make a recommendation and convince the buyer that the most profitable product matches the buyer's interest. The buyer's objective is to assess whether or not the seller is honest. The random variation in the alignment of interests allows us to infer the ability of buyers to detect deceit.

We consider two treatment variations. Our first treatment variation concerns the opportunity for the buyer to interrogate the seller. The literature on lie detection in psychology has focused almost exclusively on situations where observers are not given any opportunities for interrogation. Yet interaction is important in most relationships, and likely to affect the ability to detect deceit (Buller and Burgoon, 1996). Questions can reveal inconsistencies in someone's narrate, and elicit different behavioral responses by potential cheaters. The cognitive load to fabricate a consistent story is presumably higher when there is interaction, possibly resulting in increased leakages of signs of deception (DePaulo et al., 2003; Zuckerman et al., 1981). We therefore implemented a treatment where sellers can make a recommendation to buyers and buyers have an opportunity to interrogate the seller for a short period of time (Treatment "Questions"), and compare this to a treatment where buyers are not given this opportunity (Treatment "No-Questions").

Our second treatment variation has to do with the contextual details of the product. In the "Abstract" treatment, the product is simply a card that is either red or black. In the "Rich"

[3] Harvard Business Review, 2009. "Managing Across Distance in Today's Economic Climate: The Value of Face-to-Face Communication."

treatment, the product is a holiday package. Sellers were provided with brief descriptions and pictures of two holiday packages, one labeled red and the other labeled black. Apart from providing context, the structure of the game was identical to the Abstract treatment. The relevance of this treatment variation lies in the fact that a richer context may provide buyers with more opportunities to detect false statements.

Our findings provide support for an ability to detect deceit. Buyers are more likely to follow the seller's recommendation when it is in their interest to do so. The effect is smallest for the treatment where buyers do not have an option to interrogate the sellers. In those treatments, buyers are between seven and 11 percentage points more likely to follow the sellers recommendation when this is beneficial for them. If buyers were just randomly guessing, the expected effect would be zero. In the treatments were buyers can interrogate the sellers, the effects are substantial: 18 percentage points in the Abstract treatment and 29 percentage points in the Rich treatment.

In light of the evidence from the psychology literature, our findings are perhaps somewhat surprising (see the related literature section). There is a consensus in that literature that laymen are poor lie detectors. A main difference is that we allow for interrogation. From an evolutionary point it is not clear what we should expect about people's abilities to detect deceit. The ability to deceive is evolutionarily advantageous (Dawkins and Krebs, 1978; Wright, 1995). And indeed, deception is widespread in nature. At the same time, natural selection will favor individuals that have the ability to accurately spot attempts of deceit (Dawkins, 1978; Trivers, 1985). Strategies to deceive others and to spot deception by others are ever evolving into more subtle and effective ways, and the human brain may even have evolved accordingly (Cosmides and Tooby, 1992).

The rest of the paper is organized as follows. In the next section we discuss the related literature. We describe the experimental design in section 3. In section 4, we present the results. Section 5 concludes.

2. Related Literature

Empirical research shows that there are reliable verbal and nonverbal cues of deception, such as pupil dilation (Wang et al., 2010), type of smile (Ekman et al., 1988; ten Brinke et al., 2012), and

high-pitch voice (Ekman, 1988). Yet the consensus in the psychology literature is that untrained people without special equipment are poor lie detectors. The accuracy of deception rarely exceeds chance levels by an impressive margin, and only a small minority of studies finds an accuracy of at least ten percentage points above chance (Bond and Depaulo, 2008; DePaulo et al., 1985; Ekman and O'Sullivan, 1991; Vrij, 2008).

Several factors may have contributed to the findings in psychology. A typical drawback is that they do not allow for any interaction between potential deceivers and those who try to spot deception, or the interaction is at least partially based on predetermined transcripts (Hartwig et al., 2004). We conjecture that this may have impeded people's ability to detect deceit, and our results provide support for this view. Most of these studies also do not provide incentives for observers to accurately spot deception, nor for the potential deceivers to successfully mislead observers, although there are some exceptions such as the studies by Frank and Ekman (1997) and Kraut and Poe (1980). Another limitation of these studies is that people are commonly instructed to tell the truth or a lie. This may create a bias in the accuracy of lie detection, as poor liars that would not normally attempt to deceive anyone are now asked to deceive, and people may feel less morally burdened if they are instructed to lie (see Belot et al, 2012, for a discussion of these and other limitations). Finally, many of these studies are focused on settings that have little to do with economically relevant situations.

There are only few studies in economics that address the ability of people to detect deceit in face-to-face interactions. In some studies participants are allowed to communicate in a prisoner's dilemma, and participants or observers are asked to make predictions regarding the behavior of others (Brosig, 2002; Dawes et al., 1977; Frank et al., 1993). They find evidence that predictions are somewhat above chance levels. Belot et al. (2012) also find evidence that the accuracy of predictions by observers are above chance levels when they make predictions about contestants' choices on a game show. The contestants have to decide simultaneously to share or grab the prize money. Observers rely on informative cues such as the contestant's gender and promises. Observers cannot distinguish true promises from lies when the promise was volunteered by the contestant, but they can, to some extent, when the promise was made after the game show host asked the contestant if (s)he would share. Ockenfels and Selten (2000) asked subjects to predict if participants in a bargaining experiment had low or high bargaining costs, where costs were

randomly assigned to participants. The accuracy of predictions was above chance levels, which could be explained by objective features such as the length of the negotiations.

Some other interesting studies analyze the accuracy of predictions when participants can send free-form written messages to each other. Chen and Houser (2013) analyze messages in what they call the "mistress game." Interestingly, they find that messages contain some reliable cues of trustworthiness, such as the number of words and the mentioning of money, and they find that participants use those cues but not always in the correct manner. Utikal (2013) finds that participants can to some extent determine from a written apology whether or not an unfavorable action by the other person was intentional or accidental.

Finally, there is also a line of research that examines if communication *per se* affects behavior. Most of the experiments with face-to-face communication study behavior in a prisoner's dilemma (see Sally, 1995). Written communication is studied in a variety of other games, including coordination games, cheap talk games, and hold-up problems (e.g., Charness and Dufwenberg, 2011; Cooper et al., 1992a; Ellingsen and Johannesson, 2004a, 2004b, amongst others). In those experiments, interactions are anonymous and the message space is sometimes (but not always) restricted to a limited number of possible messages. The objective of those studies is to examine if certain types of messages affect aggregate behavior, rather than examining if people can discriminate between the sincerity or dishonesty of a particular person based on cues in that person's message or behavior.

Our game is different from the above studies, and to our knowledge we are the first to study the effects of the format of interaction (varying the opportunity for questions) and contextual richness upon the accuracy of predictions.

3. Experimental design and procedures

In our experiment, participants were matched into pairs of buyers and sellers. Sellers randomly drew a card from a deck that contained 5 red cards and 5 black cards. This draw determined if the red product or the black product was in the buyer's interest to purchase. The color of the card remained private information to the seller, and was not revealed to the buyer. In all treatments, there was a stage of 10 seconds face-to-face interaction during which sellers made a

recommendation to the buyer to "purchase" the red product or the black product. The product (i.e., the card) never physically changed hands, but instead buyers were asked to write down in private if they wanted to purchase the red or the black product.

The payoff structure for the buyer and seller was common knowledge (see Table 1). The seller earned €20 if the buyer opted for the red product, independent of the card of the seller. The buyer earned €20 if she opted for the product that matched the color of the seller's card. Thus, in terms of monetary payoffs, sellers are always better off if they can convince the buyers that they drew a red card, while the buyers are better off guessing the actual color of the seller's card. The structure of this game reflects an important class of situations, in which it is common knowledge that sellers make higher profits by selling a particular brand, but they are also the only ones to know which product is truly in the best interest of the buyers.

Table 1: Payoff Matrix

		Buyer's choice	
		Red	Black
Seller's card (random draw)	Red	20,20	0,0
	Black	20,0	0,20

Treatments

The above description was common to all treatments. In the "**No-Questions**" treatments, the only interaction time was the 10 seconds in which sellers made their recommendations. In the "**Questions**" treatments, we added another 90 seconds for the buyer to interrogate the seller. The communication in these 90 seconds was free format. In the "**Abstract**" treatment, the product description was simply a card that is either red or black. In the "**Rich**" treatment, the product is a holiday package. Sellers were provided with brief descriptions and pictures of two holiday packages, one labeled red and the other labeled black. One of these holiday packages was clearly better than the other (see the Appendix with instructions for an example). The color of the card drawn by the seller determined which holiday package was labeled as the "red holiday package" and which as the "black holiday package." Apart from providing context, the structure of the game was identical to the Abstract treatment. Sellers are always better off selling the red holiday

package, while the buyer's best interest was determined by a random draw that is private information to the seller. We provided each seller with a different set of holiday descriptions.

We used a 2x2 design giving us four different treatments. We label the four treatments as **NQ-A** (for **N**o-**Q**uestions, **A**bstract), **NQ-R**, **Q-A**, and **Q-R**. Every participant participated in only one of the context conditions (Abstract or Rich), but they all played in both interaction conditions (Questions and No-Questions). The order of the interaction condition was reversed between sessions.

Procedures

Each session consisted of 20 rounds: 10 rounds in the No-Questions treatment and 10 rounds in the Questions treatment. In every round, a seller was re-matched to a new buyer. Sellers kept the same card for 10 rounds. Participants did not receive feedback until the end of the experiment, after which one round was randomly selected for payment. We did not reveal which round was selected for payment, to preserve the anonymity of participants' decisions. The entire setup and procedures were made common knowledge to the participants, except that we did not announce the treatment variation within a session on beforehand, but only announced that there would be a second part.

For each round we collected information about the recommendation made by the seller, the purchasing decision of the buyer, the confidence by sellers that the buyer would follow their recommendation, and the confidence by buyers that the seller drew a red card. The confidence statements were not incentivized. At the end of the experiment we collected some background information as part of a survey.

The main experiment took place in Amsterdam between January and March 2012 and in October 2012, with a total of 156 participants divided over 8 sessions (46 percent female, mean age 22).[4] In two sessions we had 18 participants due to lower show up. All of the other sessions had 20 participants. Participants were students recruited from the CREED database.

[4] Prior to this experiment, we also ran four sessions with 78 participants as part of a class in experimental economics (January 2012). Participation was compulsory for those students. The incentives were the same, but they did not receive a show-up fee. Because this is a different subject pool (where people are more likely to know each other) we do not include them in the results reported here. Most of our results are robust to including those observations. We will indicate where the results differ in an important way.

Upon entering the room, participants were randomly assigned to their role as seller or buyer and they kept their role throughout the entire experiment. The instructions were distributed and read aloud by the experimenter. All participants received the same set of instructions (see Appendix A). Participants were told that the experiment consisted of two different parts, and they only received instructions for the second part after the first part had ended. Before the start of each part, sellers blindly drew a card from a deck with five red and five black cards. After showing the card to the experimenter, the experimenter would put the card back in the deck, shuffle the deck, and proceed to the next seller. The experimenter made note of the color that was drawn. The instructions explaining the game were framed in terms of a market, using terminology such as buying and selling. The descriptions of holiday packages were taken from a website (thomascook.com) and then slightly modified (an example is provided in the instructions).

The experiment was run using pencil and paper. The beginning and end of the rounds were announced by the sound of a bell. During the interactions, participants were asked to stand up. After every round, all sellers remain seated, and all buyers rotated in such a way that every buyer met every seller exactly once in each part. All participants were asked to keep their decisions sheets private. At the end of the experiment, one round was randomly selected for payment. We did not reveal which round was selected to ensure that participants could not identify the decision of any particular other participant, and participants were informed about this on beforehand. Everybody received their earnings privately in an envelope. At the end of the experiment we also administered a short questionnaire, after all decisions had been made.

The experiment lasted for about 75 minutes. Average earnings were €18.80 including a fixed show up fee of €5.

4. Results

Recommendations. We did not instruct sellers to lie or tell the truth, but with the stark incentives provided, we find that most sellers recommend the red product, even it is against the interest of the buyer. Sellers with a red card recommend the red product almost 100 percent of the time, as expected. Sellers with a black card recommend to buy the black product in less than 15 percent of the cases across all treatments.

Figure 1 shows the percentage of buyers following the seller's recommendation by treatment and recommended color. Averaged over all treatments, 83 percent of buyers follow the seller's recommendation if the recommendation is black, against 58 percent if the seller recommends red. Buyers understand that a recommendation to buy black is very likely to be truthful, and are significantly more likely to follow the recommendation in three of the treatments (Wilcoxon rank sum test, two-tailed, $Z = 1.887$, $p = 0.059$ for treatment NQ-A; $Z = 2.347$, $p = 0.019$ for NQ-R; $Z = 0.741$, $p = 0.459$ for Q-A; $Z = 3.748$, $p < 0.001$ for Q-R).

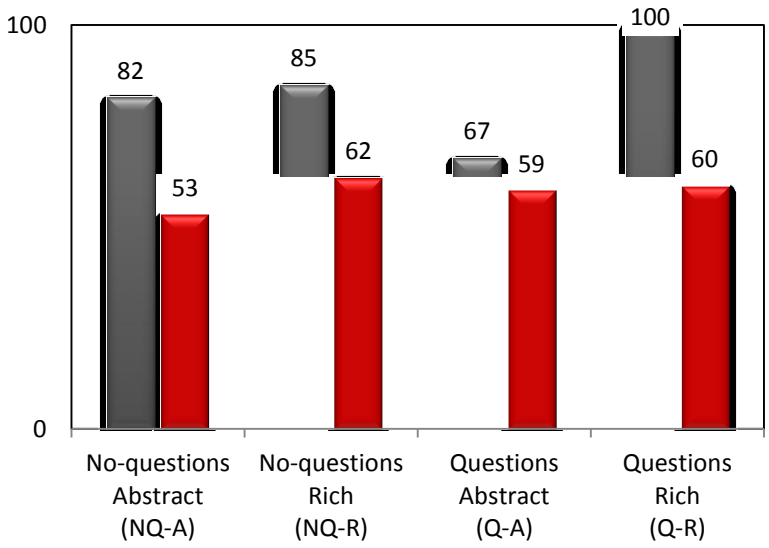

Figure 1: Percentage of buyers following the recommendation by treatment and recommended color.

Detecting deceit. The main question here is whether or not the accuracy of buyers' predictions is above chance levels. Our strategy to determine if buyers can spot deception is to compare the proportion of times that a buyer buys the red product when the seller has a red card, P(R|R), to the proportion of times that a buyer buys the red product when the seller has a black card, P(R|B). Our measure of the ability to detect deceit is thus given by:

$$D = P(R|R) - P(R|B). \quad (1)$$

If a buyer cannot identify signals of deceit, then the proportion of times that she buys the red product is independent of the card drawn by the seller, and consequently D will be 0. By contrast, D is positive if sellers are leaking signals of deception, and equals 1 if buyers can

perfectly discriminate honest from deceitful sellers. A negative value of D indicates a worse than chance accuracy.[5]

Note that our measure D is insensitive to the proportion of sellers with a red card and the proportion of times that a buyer buys the red product, in the sense that random guessing by buyers results in $D = 0$ while any positive or negative D indicates non-random guessing. Thus, some buyers may be more inclined to buy red than black if they care about the seller's payoff, but this will not lead to a positive value of D if they are guessing randomly. There might be a downward bias in D, however, if there are buyers who can guess better than chance but who decide to buy the red product whenever they are in doubt. If anything, we may therefore underestimate buyers' ability to detect deceit.

Remark. We noted before that some sellers recommend the black product. Since there is no reason for buyers to assume they are lying, it is easy for buyers to judge the honesty of the recommendation. Therefore, as a more stringent test, we exclude sellers who recommend black in what follows.

We find strong support for the idea that interaction and contextual richness are of key importance for detecting deceit. Table 2 shows the proportion of buyers that buys the red card, depending on whether the seller has a black or a red card. We also report the p-values for the test that $D = 0$ (Wilcoxon signed rank tests over the mean choices of participants[6], two tailed). In treatment NQ-A, the accuracy of detection is 7 percentage points (i.e., $D = 0.07$) and significantly different from zero. In treatment NQ-R, the ability to detect deceit is somewhat higher ($D = 0.11$) and significant. The highest accuracy is achieved in the treatments with questions, where buyers are substantially more likely to buy the red card when it is in their best interest: D=0.18 in treatment Q-A, and $D = 0.29$ in treatment Q-R, and both are significantly different from zero.[7] Figure 2

[5] While other studies have used other measures, such as the percentage of correct choices, the advantage of our measure is that it is not sensitive to the frequency with which the seller has a red card. If, for instance, in a particular session 60 percent of the sellers have a red card, then simply always buying a red card will mechanically lead to correct choices in 60 percent of the time, while accuracy would be at chance level according to the measure we use ($D = 0$).

[6] The reported nonparametric tests related to the accuracy of detecting deceit are based on taking the participant's mean over all rounds as the independent unit of observation.

[7] The results differ somewhat if we include the sessions with participants from the experimental economics class (see footnote 4). In that case, the corresponding values for D are: 0.03 in NQ-A (p =0.282), 0.06 in NQ-R (p = 0.051), 0.19 in Q-A ($p < 0.001$), 0.20 in Q-R ($p < 0.001$). We do not have enough observations to make a reliable comparison between the subject pools. We conjecture that two main factors may explain why the results are

illustrates these treatment effects at the session level. The solid circles represent the mean accuracy of buyers for each of the different sessions. To give an economic interpretation to these effects, it means that in the Question treatments, sellers with a black card can expect to make 20 to 30 percent fewer profitable sales than sellers with a red card.

Table 2: Proportion of buyers buying red (by treatment).

Treatment:	No questions		Questions	
	Abstract	Rich	Abstract	Rich
Seller has black card, P(R|B)	0.50 (0.04)	0.56 (0.04)	0.52 (0.05)	0.41 (0.05)
Seller has red card, P(R|R)	0.57 (0.04)	0.67 (0.04)	0.70 (0.05)	0.70 (0.04)
Accuracy of detecting deceit (D)	**0.07**	**0.11**	**0.18**	**0.29**
Test $D = 0$ (p-value)	0.060	0.017	0.008	0.001

Notes: Sample is sellers recommending red. Accuracy of detecting deceit measured as D = P(R|R) - P(R|B). Standard errors in parentheses. p-values of test $D = 0$ based on two-tailed Wilcoxon signed-rank tests with the mean of a participant over all rounds as the independent unit of observation.

different. First, participation for students of the experimental economics class was compulsory rather than voluntary. Second, students of the experimental economics class are more familiar with each other, as many of them take the same study program and they are more likely to have interacted with each other prior to the experiment. Alternatively, the difference is due to noise.

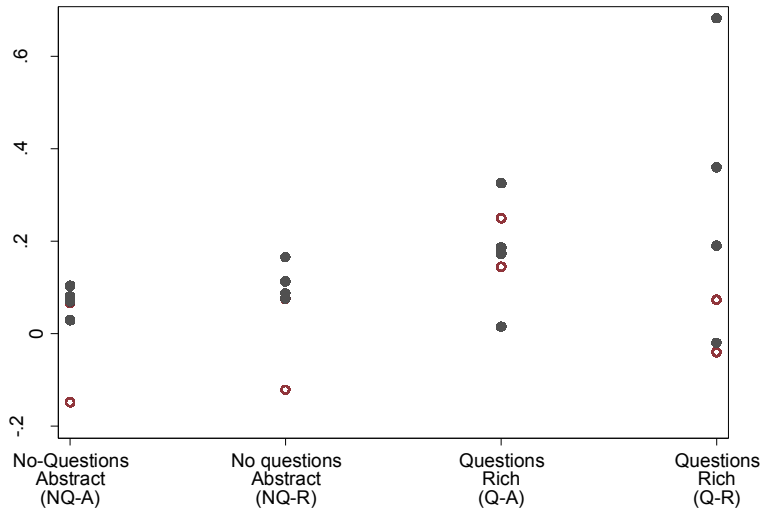

Figure 2: Mean accuracy (*D*) of buyers across treatments and sessions. The solid black circles are for participants from the CREED database, the open red circles are for participants from the Experimental Economics Class (see footnotes 4 and 7).

A regression analysis based on individual choice data confirms these results. Table 3 reports regressions where the dependent variable is the buyers' decision to choose the red product. The independent variables of interest are the interactions of the treatment and whether the seller has a red card (Treatment x red card seller). We also included treatment dummies as controls. We estimate a linear model with buyer and seller random effects. If a coefficient of the interaction term is positive, it shows that in that treatment a buyer is more likely to buy the red product if the seller has a red card compared to when the seller has a black card, and therefore if the buyer can detect deceit in a particular treatment. This is similar to the measure *D*.

We find that buyers are significantly more likely to buy the red product when the seller's card is red in all treatments except for treatment NQ-A (see column 1). For instance, in the treatment Q-R, buyers are 27.7 percentage points more likely to buy the red card if the seller has a red card, similar to what we found before. The tests reported at the bottom of the table show that the accuracy in treatment Q-R is significantly higher than in any of the other treatments. None of the other treatment effects differ significantly from each other.

Table 3: Dependent variable: Buyer buys red

Sample:	All rounds	Rounds 6-10	Confident buyers
	(1)	(2)	(3)
Treatment No-questions & Abstract x Seller has red card (1)	0.073	0.060	0.031
	(0.049)	(0.072)	(0.075)
Treatment No-questions & Rich x Seller has red card (2)	0.113**	0.139*	0.035
	(0.051)	(0.073)	(0.078)
Treatment Questions & Abstract x Seller has red card (3)	0.120**	0.023	0.251***
	(0.052)	(0.074)	(0.074)
Treatment Questions & Rich x Seller has red card (4)	0.277***	0.276***	0.307***
	(0.051)	(0.073)	(0.069)
Tests equality of coefficients (p –values)			
(2) = (1)	0.574	0.446	0.968
(3) = (1)	0.516	0.771	0.035
(3) = (2)	0.924	0.299	0.045
(4) = (3)	0.031	0.019	0.578
Controls	Yes	Yes	Yes
Observations	1,431	704	605
Number of groups	78	78	71

Notes: Two-way linear random effects model (allowing for buyer and seller random effects). Sample is sellers recommending the red product. Control variables: Treatment dummies. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Learning. A natural question is if learning matters. The scope for learning is somewhat restricted in our experiment, because we do not give any feedback to participants. In the treatments without interrogation we do not expect to see learning for that reason. But conceivably some learning can take place in the Question treatments: with experience, sellers may become more convincing in their recommendations, and buyers may learn to ask more relevant questions. Thus, a priori, the ability to detect deceit could increase or decrease over time. Column (2) of Table 3 shows the estimates for the second half of the treatment. The results are similar, except that in treatment Q-A we no longer find evidence of buyers detecting deceit. On net, there seems to be little learning.

Confidence. Buyers vary in their degree of confidence when they meet different sellers. In column (3) of Table 3 we restrict the sample to buyers that report to be confident in their choice.

We categorize a choice as confident if the reported confidence is above the median reported confidence. In the No-Questions treatments, we do not find a higher accuracy among confident buyers. If anything, the accuracy of confident buyers is lower rather than higher. In the Questions treatment, we find that the accuracy is higher in the Abstract condition, but we find little difference in the Rich condition.

Background characteristics. Table 4 shows the results of a regression of a correct decision by a buyer (i.e., a buyer that buys the color matching that of the seller) on several background characteristics of the participants. Background characteristics are organized by the role of participants. If certain characteristics are positively (/negatively) related to the ability to detect deceit, then we should expect the coefficients of these variables to be positive (/negative) for buyers. If certain background characteristics are positively (/negatively) related to the ability to deceive, we should expect the coefficients of these variables to be negative (/positive) for sellers.

The background characteristics we include are gender and a dummy variable that equals 1 if the buyer-seller pair in a particular round consists of two females. We also include a self-reported measure of the participant's perceived own ability to deceive or spot deception (both on a seven point Likert scale, ranging from 1: not good at all, to 7: very good) and how often they play poker (on a five scale ranging from 1: never, to 5: once or twice a day).

We report the results separately for sellers with a black card and sellers with a red card. Columns (1) and (2) of Table 4 report regressions for all treatments combined. Because we found no statistical difference in accuracy between any treatments except for the treatment with questions and rich context, we look separately at the latter (columns 3 and 4) and the other treatments pooled (columns 5 and 6).

Consistent with results from a meta-study (Bond & Depaulo, 2008), we do not find a strong indication that any of these variables is related to the ability to deceive or spot deception. Most coefficients are insignificant and not consistent in sign across specifications. The only variable that is significant is the frequency of poker: sellers who play poker more often are more likely to deceive buyers when they have a black card.

Table 4: Accuracy of choices and background characteristics

Dep. variable: correct prediction	(1)	(2)	(3)	(4)	(5)	(6)
Treatments:	All treatments		Questions and rich context		Other treatments	
Seller's card:	Black card	Red card	Black card	Red card	Black card	Red card
Buyer characteristics						
Female	0.003	-0.063	0.099	-0.055	-0.023	-0.080
	(0.063)	(0.062)	(0.130)	(0.094)	(0.072)	(0.072)
Self-perceived ability to detect deceit	0.015	-0.003	0.004	-0.019	0.012	-0.002
	(0.022)	(0.022)	(0.039)	(0.035)	(0.025)	(0.024)
Frequency poker	0.025	-0.012	0.073	0.045	0.022	-0.021
	(0.037)	(0.037)	(0.065)	(0.061)	(0.042)	(0.040)
	0.003	-0.063	0.099	-0.055	-0.023	-0.080
Seller characteristics						
Female	0.051	-0.057	0.010	-0.011	0.055	-0.075
	(0.054)	(0.049)	(0.128)	(0.096)	(0.061)	(0.058)
Self-perceived ability to deceive	-0.008	-0.001	-0.018	-0.013	-0.003	0.003
	(0.011)	(0.011)	(0.026)	(0.022)	(0.013)	(0.013)
Frequency poker	-0.084***	-0.033	-0.164**	0.001	-0.060**	-0.050
	(0.027)	(0.026)	(0.067)	(0.052)	(0.030)	(0.031)
Joint characteristics						
Both female	-0.027	0.105	-0.112	0.168	-0.000	0.082
	(0.076)	(0.070)	(0.164)	(0.120)	(0.087)	(0.088)
Observations	686	730	143	213	543	517
Groups	77	77	39	39	77	77

Two-way linear random effects model (allowing for buyer and seller random effects). Sample is sellers recommending the red product. Columns 5 and 6: All treatments except the treatment Questions and rich-context. Standard errors in parentheses.

*** p<0.01, ** p<0.05, * p<0.1.

5. Discussion

Our results are relevant for a wide range of economic exchange situations. Economists have extensively studied markets with asymmetric information, in which some parties are better informed and have objectives that are not fully aligned with those of the other parties. Salesmen, financial advisors, and manufacturers of products are all better informed about the products they sell than their vulnerable customers. The predominant view is that this causes inefficiencies in, or even nonexistence of, markets (Akerlof, 1970). Our results show why these inefficiencies can be limited and explain why face-to-face interactions are highly valued in business (Harvard Business Review, 2009)*:* people do not always buy lies if they can interrogate.

We close our paper by mentioning some limitations to our setup. Although we have established the importance of some factors in the ability to spot deception, we are uncertain about the cues that buyers use. Answering that question would require experiments where conversations are recorded and categorized (as for instance in Brandts et al., 2012), and where different communication channels are studied in isolation. One key issue of course is that such recording could affect the behavior of participants in a non-trivial way. Another interesting question is whether buyers and sellers with certain characteristics will self-select on markets. Perhaps in real markets sellers are relatively effective at deceiving people, or perhaps people will only buy a product if they can trust the seller. Furthermore, in our experiment buyers are aware of the possibility of deception, whereas in real life they may be more credulous (Irlenbusch and Ter Meer, 2013). Another aspect of our design is that the private information of sellers concerns the state of nature, rather than their own actions. Serra-Garcia et al. (2013) show that people are less likely to lie about their actions than about their knowledge of the state of nature. Clearly, many extensions are possible, and we hope that our framework provides an avenue for future research. Given the limited knowledge that is currently available, we believe that this is an important topic that deserves more study.

APPENDIX A: Instructions (Not for publication)

Below are the instructions for the no-context treatment that starting with the no-questions treatment. Instructions for the sessions with the interaction treatment first were very similar with some obvious changes.

Instructions for the no-context treatment.

Welcome to this experiment. Please read the instructions carefully. We ask you to remain silent throughout the experiment unless indicated otherwise.

There are two types of participants in this experiment: "sellers" and "buyers." Once you have been randomly assigned a type, you will keep that same type for the entire experiment. The experiment consists of two parts, each with 10 rounds. Sellers and buyers are re-matched every round such that no two participants will ever be paired more than once with each other within each part of the experiment.

One of the rounds of the two parts will be randomly selected at the end of the experiment, and you will be paid your earnings of that round. We will not make public which round has been selected for payment, to ensure that you cannot identify your partner of that round and vice versa. You will also receive €5 as a fixed show up fee.

Part 1.

What follows is a description of the tasks for both types for the first 10 rounds of the experiment.

In the first round, seller 1 is matched to buyer 1, seller 2 to buyer 2, etc. The seller sits to the right of the buyer when facing the computer screen.

Sellers: sellers draw (without looking) a card from a deck of 10 cards, of which five are red and five are black. After seeing the card, the seller will give the card back, we will put it back in the deck and proceed to the next seller. The seller will keep the same color of the card for all 10 rounds. Every seller is paired with one buyer. The seller must make a recommendation to the buyer to “buy” a red card or a black card.

Buyers: The buyer does not observe which card is drawn by the seller. After the seller has made a recommendation, the buyer must decide to buy the red card or the black card.

Recommendation phase: The seller has 10 seconds to make a recommendation. The starting time and ending time are announced by the sound of a bell. Only the seller is allowed to speak during that time. Please do not speak too loudly.

During this recommendation phase, all pairs of participants stand up and face each other, as in the diagram below. The rest of the time you remain seated and silent.

buyer seller buyer seller

After the recommendation phase, sellers and buyers fill in the decision forms *in private*.

The seller writes down:

1. The number of the buyer,
2. The recommendation he or she made (black or red),
3. How certain (s)he feels that the buyer buys a black or red card (on a 10 point scale).

The buyer writes down:

1. The number of the seller,
2. The card he or she wishes to buy (black or red),
3. Which color the seller recommended,
4. How certain (s)he feels that the color of the card that the seller drew is black or red (on a 10 point scale).

Earnings: The seller earns €20 if the red card is bought by the buyer, independent of whether or not the seller actually has a red or a black card. The buyer earns €20 if he bought the card with the same color that the seller has actually drawn. Thus, the seller always earns more if the buyer buys a red card. The buyer always earns more if he buys a card of the same color as the seller drew. There are four possibilities:

1. The seller has a red card. The buyer buys the red card. Both earn €20.
2. The seller has a red card. The buyer buys the black card. Both earn nothing.
3. The seller has a black card. The buyer buys the red card. The seller earns €20, the buyer nothing.
4. The seller has a black card. The buyer buys the black card. The seller earns nothing, the buyer earns €20.

After completing the form for a round, please wait until we announce that we go to the next round. All sellers remain seated, and all buyers move to the next seller. Sellers are numbered 1 through 10. So if you are matched with seller 1 in the first round, you would then go to seller 2. After being paired with seller 10 you go to seller 1.

Here is a summary of the steps:

1. The seller draws a card,
2. When the bell rings each pair stands up facing each other, and the seller makes a recommendation during 10 seconds,

3. When the bell rings again you sit down and fill in the form in private for that round,
4. Buyers move to the next seller and we repeat steps 2 to 4.

Please always keep your sheets private so that others cannot see your decisions.

If you have any questions, raise your hand and wait for the experimenter.

Part 2.

What follows is a description of the tasks for both types for the next (and last) 10 rounds of the experiment.

The second part proceeds in a very similar way as the first part. The only difference is that after the recommendation phase of 10 seconds (announced by the sound of the bell), there will be another 90 seconds in which the buyer can ask questions to the seller and the seller can respond. The communication in these 90 seconds is completely free as long as no offensive, threatening, or indecent language is used. You remain standing facing your paired participant until we announce that the 90 seconds have passed at which point you sit down and complete the decision forms *in private*.

You start again in the same position as in part 1. All sellers will draw another card from the deck.

If you have any questions, raise your hand and wait for the experimenter.

Instructions for rich-context treatment.

The instructions for the rich-context treatment were the same except for the following:

a) Buyers were told they could buy a red or black holiday package (instead of card)

b) The part describing the task for sellers had the following additional text:

> Sellers then receive a sheet with a description of two holiday packages. The two holiday packages have the same destination, but one package is more attractive for the buyer than the other. The best package for the buyer will get the label of the card drawn from the deck by the seller. Thus, if the seller draws a red card, the better value package for the buyer is identified as "package Red," while if the seller drew a black card, the better value package for the buyer is identified as "package Black." The seller will keep the same sheet for all 10 rounds. The holiday destination is different for each seller.
>
> Sellers will get 5 minutes to study the sheet with the descriptions of holiday packages. They can keep this sheet with them throughout the part. An example is given at the end of these instructions.

c) The part on earnings was replaced by:

Earnings: The seller earns €20 if the red package is bought by the buyer, independent of whether or not the seller actually drew the red or the black card. The buyer earns €20 if (s)he bought the package that has the best value for the buyer. Thus, the seller always earns more if the buyer buys the red package. The buyer always earns more if he buys the best package (which is the red package if the seller drew a red card, and the black package if the seller drew a black card).

There are four possibilities:

1. The seller drew a red card (so the red package has the best value for buyer). The buyer buys the red package. Both earn €20.
2. The seller drew a red card (so the red package has the best value for buyer). The buyer buys the black package. Both earn nothing.
3. The seller drew a black card (so the black package has the best value for buyer). The buyer buys the red package. The seller earns €20, the buyer nothing.
4. The seller drew a black card (so the black package has the best value for buyer). The buyer buys the black package. The seller earns nothing, the buyer earns €20.

d) The following example of the holiday packages was provided to all participants. For every seller and every part there was a different set of holiday packages.

OPTION A: LOW VALUE

Kervansaray Lara Resort
Lara Beach, Turkey
***

Key Features

- Two swimming pools
- sun terraces with sunloungers and parasols
- indoor pool
- buffet restaurant

The Kervansaray Lara Resort is modern in style and is situated on the beachfront. This hotel has amazing architecture and good facilities and will particularly appeal to couples and families who wish to enjoy a beach holiday in a quiet resort.

Average Customer Ratings

- Holiday overall 71%
- Accommodation 69%
- Location 66%
- Food 70%

REMARKS:

food in hotel often cold, far from the beach, no good connection to airport

OPTION B: HIGH VALUE

Alba Royal
Nr Side, Turkey
*****

Key Features

- Outdoor free–form swimming pool and sun terraces with sunloungers and parasols (at least one pool open Nov–Mar)
- buffet restaurant
- three à la carte restaurants
- 24hr room service (not part of All Inclusive)

Situated directly on the sandy beach with stunning views of the sea, the stylish Alba Royal offers a host of sporting activities, and is ideal for couples. This hotel offers bright, comfortable rooms and an excellent selection of facilities, including a choice of bars and restaurants and a modern, well–equipped spa.

Average Customer Ratings

- Holiday overall 91%
- Accommodation 92%
- Location 88%
- Food 90%

Akerlof, G.A., 1970. The Market for "Lemons": Quality Uncertainty and the Market Mechanism. Q. J. Econ. 84, 488–500.

Belot, M., Bhaskar, V., van de Ven, J., 2012. Can observers predict trustworthiness? Rev. Econ. Stat. 94, 246–259.

Bond, C.F., Depaulo, B.M., 2008. Individual differences in judging deception: accuracy and bias. Psychol. Bull. 134, 477–92.

Brandts, J., Charness, G., Ellman, M., 2012. Let's talk: How communication affects contract design [WWW Document]. Mimeo.

Brosig, J., 2002. Identifying cooperative behavior: some experimental results in a prisoner's dilemma game. J. Econ. Behav. Organ. Organ. 47, 275–290.

Buller, D.B., Burgoon, J.K., 1996. Interpersonal Deception Theory. Commun. Theory 6, 203–242.

Charness, G., Dufwenberg, M., 2006. Promises and Partnership. Econometrica 74, 1579–1601.

Charness, G., Dufwenberg, M., 2011. Participation. Am. Econ. Rev. 101, 1211–1237.

Cooper, R., DeJong, D. V., Forsythe, R., Ross, T.W., 1992. Communication in Coordination Games. Q. J. Econ. 107, 739–771.

Cosmides, L., Tooby, J., 1992. Cognitive adaptations for social exchange, in: Barkow, J., Cosmides, L., Tooby, J. (Eds.), The Adapted Mind : Evolutionary Psychology and the Generation of Culture. Oxford University Press, Oxford.

Dawes, R.M., Mctavish, J., Shaklee, H., 1977. Behavior, communication, and assumptions about other people's behavior in a commons dilemma situation. J. Pers. Soc. Psychol. 35, 1–11.

Dawkins, R., 1978. The selfish gene. New York, NY : Oxford University Press, New York, NY.

Dawkins, R., Krebs, J.R., 1978. Animal signals: Information or manipulation'?, in: Krebs, J.R., Davies, N.B. (Eds.), Behavioral Ecology. Blackwell Publishers, Oxford.

DePaulo, B.M., Lindsay, J.J., Malone, B.E., Muhlenbruck, L., Charlton, K., Cooper, H., 2003. Cues to deception. Psychol. Bull. 129, 74–118.

DePaulo, B.M., Stone, J., Lassiter, D., 1985. Deceiving and detecting deceit, in: Schenkler, B. (Ed.), The Self and Social Life. McGraw-Hill, New York, NY.

Ekman, P., 1988. Lying and nonverbal behavior: Theoretical issues and new findings. J. Nonverbal Behav. 12, 163–175.

Ekman, P., Friesen, W. V., O'Sullivan, M., 1988. Smiles when lying. J. Personal. Soc. Psychol. 54, 414–420.

Ekman, P., O'Sullivan, M., 1991. Who can catch a liar? Am. Psychol. 46, 913–20.

Ellingsen, T., Johannesson, M., 2004a. Is There a Hold-up Problem? Scand. J. Econ. 106, 475–494.

Ellingsen, T., Johannesson, M., 2004b. Promises, Threats and Fairness*. Econ. J. 114, 397–420.

Frank, M.G., Ekman, P., 1997. The ability to detect deceit generalizes across different types of high-stake lies. J. Pers. Soc. Psychol. 72, 1429–39.

Frank, R.H., Gilovich, T., Regan, D.T., 1993. The evolution of one-shot cooperation: An experiment. Ethol. Sociobiol. 14, 247–256.

Gneezy, U., 2005. Deception: The Role of Consequences. Am. Econ. Rev. 95, 384–394.

Hartwig, M., Granhag, P.A., Strömwall, L.A., Vrij, A., 2004. Police Officers' Lie Detection Accuracy: Interrogating Freely Versus Observing Video. Police Q. 7, 429–456.

HBR, 2009. Managing across distance in today's economic climate: The value of face-to-face communication. Harv. Bus. Rev. 1–15.

Irlenbusch, B., Ter Meer, J., 2013. Fooling the Nice Guys: Explaining receiver credulity in a public good game with lying and punishment. J. Econ. Behav. Organ. 93, 321–327.

Kartik, N., 2009. Strategic Communication with Lying Costs. Rev. Econ. Stud. 76, 1359–1395.

Kartik, N., Ottaviani, M., Squintani, F., 2007. Credulity, lies, and costly talk. J. Econ. Theory 134, 93–116.

Kraut, R.E., Poe, D.B., 1980. Behavioral roots of person perception: The deception judgments of customs inspectors and laymen. J. Pers. Soc. Psychol. 39, 784–798.

Sally, D., 1995. Conversation and Cooperation in Social Dilemmas. Ration. Soc. 7, 58–92.

Serra-Garcia, M., van Damme, E., Potters, J., 2013. Lying about what you know or what you do? J. Eur. Econ. Assoc. 11, 1204–1229.

Ten Brinke, L., Porter, S., Baker, A., 2012. Darwin the detective: Observable facial muscle contractions reveal emotional high-stakes lies. Evol. Hum. Behav. 33, 411–416.

Trivers, R., 1985. Social Evolution. Benjamin/Cummings, Menlo Park, CA.

Utikal, V., 2013. I am sorry - Honest and fake apologies.

Vrij, A., 2008. Detecting Lies and Deceit: Pitfalls and Opportunities. Wiley, Chichester.

Wang, B.J.T., Spezio, M., Camerer, C.F., 2010. Pinocchio ' s Pupil : Using Eyetracking and Pupil Dilation to Understand Truth Telling and Deception in Sender-Receiver Games. Am. Econ. Rev. 100, 984–1007.

Wright, R., 1995. The Moral Animal: Why we are the way we Are: The new science of evolutionary psychology.

Zuckerman, M., DePaulo, B.M., Rosenthal, R., 1981. Verbal and nonverbal communication of deception, in: Advances in Experimental Social Psychology. Academic Press, New York, NY, pp. 1–57.